Measuring the average outcome and inequality effects of segregation in the presence of social spillovers[1]

Bryan S. Graham[†] Guido W. Imbens[‡] Geert Ridder[♮]

Initial Draft: September 30, 2005
This Draft: December 29, 2005

Preliminary and Incomplete

Comments/Corrections Welcome

Abstract: In this paper we provide a nonparametric treatment of identification in models with social spillovers. We consider a setting with 'high' and 'low' type individuals. Individual outcomes depend upon the fraction of high types in one's group. We refer to this dependence as a social spillover or peer group effect. We define estimands measuring local and global spillover strength as well as the outcome and inequality effects of increasing segregation (by type) across groups. We relate our estimands to the theory of sorting in the presence of social externalities.

JEL Classification: C14, C31, D62, I21

Key Words: Social Interactions, Peer Group Effects, Sorting, Segregation, Equity vs. Efficiency Trade-off.

[1] Prepared for the 2006 AEA Meetings. We would like to thank participants of the Berkeley Labor Lunch for helpful comments. All the usual disclaimers apply.

[†] Department of Economics, University of California - Berkeley, 549 Evans Hall #3880, Berkeley, CA 94720. E-mail: bgraham@econ.berkeley.edu.

[‡] Department Agricultural and Resource Economics and Department of Economics, University of California - Berkeley, 330 Giannini Hall, Berkeley, CA 94720. E-mail: imbens@econ.berkeley.edu.

[♮] Department of Economics, University of Southern California, Kaprielian Hall, University Park Campus, Los Angeles, CA 90089. E-mail: ridder@usc.edu.

1 Introduction

In this paper we provide a nonparametric treatment of identification in models with social spillovers. We consider a setting where individuals are either 'high' or 'low' types. Individual outcomes may vary with the fraction of high type individuals in their social group. We refer to this dependence as a social spillover or peer group effect. An example, which we will carry throughout the paper, is the relationship between own academic achievement and the socioeconomic background, race or gender of one's classmates.

We define estimands measuring local and global spillover strength as well as the outcome and inequality effects of increasing segregation (by type) across groups. These latter estimands are particularly novel and policy-relevant. We focus on characterizing the effects of 'local' reallocations of individuals across groups such that the marginal distribution of types in the population of interest remains unchanged. These estimands allow us to assess the efficiency of the status quo assignment of individuals to groups as well as the possibility of equity versus efficiency trade-offs associated with moves toward either increased segregation or integration.

Our framework offers two important advantages over existing approaches to learning about social spillovers. First, we define estimands that speak directly to the (albeit sometimes implicit) policy question of interest: are deviations from the status quo assignment of individuals to groups welfare-increasing? Our approach, for example, would be useful to policy-makers considering ability tracking within schools, school busing to promote desegregation, different mechanisms to encourage public housing recipients to move to more affluent neighborhoods or different approaches to organizing jails.[2]

Second, our approach explicitly connects 'the data' with many of the ideas emphasized in theoretical work on sorting in the presence of social spillovers, e.g., de Bartolome (1990), Benabou (1993), Epple and Romano (1998) and Becker and Murphy (2000).[3] The connection between empirical work on social externalities and the relevant applied public finance theory has (so far) remained limited.[4] This is problematic as the theoretical literature identifies "the exact conditions that need to be empirically estimated" in order to think coherently about the policy implications of social externalities (Piketty 2000, p. 464).

Our estimand of the average outcome effect of a local increase in segregation, for example, is related to the theoretical literature's characterization of the social planner's problem of optimally assigning individuals to groups. Another connection is determining whether the *laissez faire* allocation is efficient. Implementing, for example, 'school choice' policies or providing rental vouches to public housing recipients arguably increases the role of market forces in assigning individuals to groups. It is important to assess whether such policies would increase segregation and to characterize the equity and efficiency implications of any such increases. We provide methods for answering these and other questions raised

[2]Recent empirical research in these areas includes, among others, Epple, Newlon and Romano (2002) and Figlio and Page (2002) on ability tracking in schools, Angrist and Lang (2004), Guryan (2004) and Card and Rothstein (2005) on the effect of school desegregation on educational achievement and attainment, Kling, Leibman and Katz (2005) on the effects of neighborhood quality on a variety of outcomes, and Bayer, Pintoff and Pozen (2004) on peer effects in juvenile detention. There are also older literatures on these questions and others in sociology and education (e.g., Coleman et al 1966).

[3]Much of this theoretical literature is surveyed by Piketty (2000) and Fernandez (2003).

[4]A very important exception to this statement is a small 'structural' empirical public finance literature, e.g., Epple and Sieg (1999) and Bayer, Ferreira and McMillan (2003). Hedonic models might provide a complementary approach to learning about social spillovers (e.g., Epple 1987, Ekeland, Heckman and Nesheim 2004).

by theoretical work. The nonparametric set-up is essential to achieving these goals.

Section 2, which follows next, motivates, derives and interprets the nonparametric causal functions we seek to identify. We begin by defining local and global measures of spillover strength. We then look at the effects of a local increases in segregation on average outcomes and inequality. We also provide conditions under which a shift toward the *laissez faire* allocation of individuals to groups is likely to (a) increase or decrease segregation and (b) raise or lower average outcomes. Our approach thus provides a natural framework for coherently thinking about equity versus efficiency concerns that typically arise in policy debates. For example, would greater school choice induce 'cream skimming' and what would be its effect on average achievement and inequality? This section also develops connections between our proposed estimands and the theory of sorting in the presence of social externalities.

Section 3 relates our general nonparametric approach to specific parametric models of social interactions. In particular we draw connections with the 'linear-in-means' model which currently dominates applied work. We also examine a simple generalization of this model which addresses some of the main shortcomings of the linear-in-means model. This section also provides a transparent illustration of the key concepts introduced in section 2, using simple and familiar models as vehicles. Section 4 specifies conditions under which our estimands are identified and discusses settings where these conditions will be most compelling. Section 5 summarizes and suggests area for future research. In ongoing work we also provide a comprehensive treatment of estimation and inference issues. In this note, however, we focus primarily upon defining our estimands and identification.

2 Definitions of causal functions of interest

We observe $c = 1, \ldots, C$ social groups (e.g., classrooms or neighborhoods). In group c there are $j = 1, \ldots, N_c$ individuals for a total of $n = \sum_{c=1}^{C} N_c$ individuals overall. The 'type' of j^{th} individual in the c^{th} group is given by $T_{cj} \in \{L, H\}$. We refer to type H individuals as 'high' types and type L individuals as 'low' types. An individual's type denotes a permanent characteristics such as ethnicity or gender.[5] Also available is a vector of observed group characteristics, X_c. This vector with typically include group size as well as, say, teacher or school characteristics or attributes of a neighborhood. The composition of the c^{th} group is given by the fraction of high type individuals in the group, i.e., $S_c = \sum_{j=1}^{N_c} \mathbf{1}\left(T_{cj} = H\right)/N_c$, with $S_c \in [0,1]$. We assume that N_c is large enough such that we can treat S_c 'as if' it were a continuous random variable.

For what follows it is helpful to also establish a single index notation. Let $i = 1, \ldots, n$ index individuals with $G_i \in \{1, \ldots, C\}$ being a group indicator such that S_{G_i} is the group composition for individual i; n_H and n_L denote the number of high and low types in the sample with $n = n_H + n_L$.

In the spirit of Angrist, Imbens and Graddy (2000), we posit the existence of continuously differentiable – individually heterogeneous – response functions of the form

$$Y_i\left(s\right), \quad \text{for all } s \in [0,1]. \tag{1}$$

Equation (1) provides an *individual-specific* mapping from 'potential' group compositions into outcomes.

[5] We are currently extending our setup to the case where T_{cj} denotes an individual treatment such as a vaccination or exposure to an education program.

An individual's actual or realized outcome is given by $Y_i = Y_i(S_{G_i})$.

We now define the main causal functions we seek to identify. The key building block for all of our estimands is the conditional expectation function (CEF) of $Y(s)$ given type and group characteristics. This potential outcome CEF for individuals of type $T = t$ in groups with characteristics $X = x$ equals

$$m(s,t,x) \overset{def}{\equiv} \mathbb{E}[Y(s)|T = t, X = x]. \tag{2}$$

Equation (2) is a causal function of s, in the sense that it is identifiable through experimentation: at $T = t$ and $X = x$ experimental (i.e., random) variation in S identifies $m(s,t,x)$. For convenience we define $m_H(s,x) \overset{def}{\equiv} m(s,H,x)$ and $m_L(s,x) \overset{def}{\equiv} m(s,L,x)$. Unless stated otherwise, expectations are taken across individuals (not groups). At times we will employ the single index notation to emphasize this point.

The expected average outcome if we create of group of composition $S = s$ is given by

$$m(s,x) \overset{def}{\equiv} \mathbb{E}[Y(s)|X = x] = s \cdot m_H(s,x) + (1-s) \cdot m_L(s,x), \tag{3}$$

Average outcomes in a group equal a weighted average of the type-specific expected outcomes given by (2).

Differentiating (3) with respect to s gives the average partial effect of changes in group composition on expected average group outcomes:

$$m'(s,x) = m_H(s,x) - m_L(s,x) + s \cdot m'_H(s,x) + (1-s) \cdot m'_L(s,x). \tag{4}$$

This average partial effect consists of two parts. The first part, $p(s,x) \overset{def}{\equiv} m_H(s,x) - m_L(s,x)$, is the effect of changing group composition on expected outcomes holding spillover strength constant. It is the compositional or *private effect* of changing group composition on expected average outcomes. The second component, $e(s,x) \overset{def}{\equiv} s \cdot m'_H(s,x) + (1-s) \cdot m'_L(s,x)$, measures the spillover or *external effect* associated with increasing s.

Replacing a low type individual with a high type individual raises the group average outcome for two reasons. First, irrespective of the presence of social spillovers, average outcomes will rise because the composition of the group has shifted toward high types. This effect is private, in the sense that it captures benefits that are entirely confined to the entering high type. Second, the introduction of an additional high type individual into the group creates a spillover which raises outcomes for all individuals in the group.

2.1 Measuring spillover strength

A measure of local spillover strength at $S = s$ is provided by averaging $e(s,X)$ over the marginal distribution of X

$$\alpha^{LSS}(s) \overset{def}{\equiv} \mathbb{E}_X[e(s,X)]. \tag{5}$$

Note that $\alpha^{LSS}(s)$ only differs from zero in the presence of social spillovers. The local nature of $\alpha^{LSS}(s)$ is important for discriminating between different theories of social spillovers. In the classroom setting, for example, disruption models of peer groups effects imply that a few 'bad apples' can spoil the learning environment for all students (c.f., Lazear 2001). In such a setting we might expect $\alpha^{LSS}(s)$ to be positive and large for values of s close to one but close to zero for other group compositions (since once a few disruptive students are in a classroom the effect of adding more is marginal).

A measure of overall spillover strength is provided by averaging $e(S,X)$ over the joint distribution of S and X,

$$\alpha^{ASS} \stackrel{def}{\equiv} \mathbb{E}[e(S,X)]. \tag{6}$$

α^{ASS} provides a natural summary measure of the strength of social spillovers in the population of interest.

Previous empirical work emphasizes the concept of a social multiplier (e.g., Manski 1993, Brock and Durlauf 2001, Glaeser and Schienkman 2003). The social multiplier is the ratio of the full-to-compositional effect of changes in s on expected average outcomes. Averaging this quantity over the marginal distribution of X and the joint distribution of S and X gives, respectively, the local and average social multipliers

$$\gamma^{LSM}(s) \stackrel{def}{\equiv} 1 + \mathbb{E}_X\left[\frac{e(s,X)}{p(s,X)}\right], \qquad \gamma^{ASM} \stackrel{def}{\equiv} 1 + \mathbb{E}\left[\frac{e(S,X)}{p(S,X)}\right].$$

In the absence of social spillovers both the local and average social multipliers equal one since the effect of changes in the fraction of high types on expected average outcomes in this case is purely compositional. We focus on identifying and estimating (5) and (6) in this paper, but our basic results readily extend to $\gamma^{LSM}(s)$ and γ^{ASM}.

2.2 Measuring the efficiency and equity effects of segregating-increasing reallocations

Estimands (5) and (6) provide measures of the strength of social spillovers. From the perspective of a social planner, however, interest centers on optimally assigning individuals to different groups, holding the marginal distribution of individual types fixed. Will a shift toward more segregation (by type) raise average outcomes? Will a shift toward more integration do so?

We consider shifts in the distribution of S such that the fraction of high type agents in the c^{th} group changes according to

$$\Delta S_c = \lambda(S_c - \mu_T) \tag{7}$$

where $\mu_T = \mathbb{E}[T]$ and λ is positive but small enough to ensure that $S_c + \Delta S_c$ is below one for all groups. The distributions of group-size and X are held constant across reallocations. Importantly, (7) does not alter the marginal distribution of individual types and is therefore feasible.[6] Let σ_S equal the status quo standard deviation of S (*across individuals*), implementing (7) increases σ_S by $100 \times \lambda$ percent to $(1+\lambda)\sigma_S$; it is in this sense that (7) is 'segregation-increasing'. Segregation-decreasing reallocations can be defined by setting $\lambda \in [-1, 0)$; λ equal to -1 results in homogenous groups or perfect integration.

For a group with an initial composition of $S = s$ and $X = x$ the expected change in outcomes

[6] That (7) is feasible is straightforward to show. By the linearity of the expectations operator and iterated expectations we have that $\mathbb{E}[S_i] = \mathbb{E}[\mathbb{E}[T_i|N_i]] = \mathbb{E}[T_i]$. For $S_i^* = S_i + \Delta S_i$ we have

$$S_i^* = (1+\lambda)S_i - \lambda\mu_T,$$

and hence $\mathbb{E}[S_i^*] = \mathbb{E}[S_i] = \mathbb{E}[T_i]$. The reallocation leaves the fraction of high types in the population unchanged.

associated with implementing reallocation (7) is given by

$$r(s,x,\lambda) \stackrel{def}{\equiv} m(s+\lambda(s-\mu_T),x) - m(s,x).$$

A segregating-increasing reallocation raises (lowers) expected average outcomes in a group if $\frac{\partial r(s,x,0)}{\partial \lambda} = m'(s,x)(s-\mu_T)$ is positive (negative), where

$$m'(s,x)(s-\mu_T) = p(s,x)(s-\mu_T) + e(s,x)(s-\mu_T).$$

Expected average outcomes across all individuals therefore increase (decrease) if

$$Cov\left(m'(S,X),S\right) = Cov\left(p(S,X),S\right) + Cov\left(e(S,X),S\right) \tag{8}$$

is positive (negative).

Equation (8) is an intuitive condition. If groups where the fraction of high type individuals exceeds the population mean ($S > \mu_T$) are also relatively responsive to changes in s, then reallocations that reinforce any existing segregation will tend to raise average outcomes. In contrast, if groups with a low fraction of high types are more responsive to changes in s, then such reallocations will tend to lower average outcomes. For λ small, $\lambda \times Cov(m'(S,X),S)$ approximately equals the change in average outcomes associated with implementing (7).

Interpreting deviations of $Cov(m'(S,X),S)$ from zero requires some care. Reallocations of individuals to groups may alter population average outcomes for three distinct reasons. First, peer quality changes for those individuals who change groups as part of the reallocation. We call this the internalizable or *private peer effect*. Second, group characteristics for those individuals who switch groups also change. For reasons outlined below we call this the *targeting effect* (c.f., Piketty 2004). Finally, peer quality changes for those individuals who do not switch groups as part of the reallocation, called 'stayers', we call this the *external (peer) effect.*

The first two effects are captured by the first component of (8), while the external effect is captured by the last component. First consider the latter term, it captures changes in average outcomes operating through the reallocation's effect on average spillover strength. If the marginal benefit of an additional high type on stayers is greater in groups with a large fraction of high types (i.e., $Cov(e(S,X),S) > 0$), then increased segregation will raise average outcomes by raising average spillover strength. This term is only non-zero in the presence of some form of social spillover. The sign of $Cov(e(S,X),S)$ determines the direction of the external effect associated with implementing (7).

The first term in (8), $Cov(p(S,X),S)$, reflects a combination of the private peer and targeting effects. Both of these effects are 'private' in the sense that they capture effects that operate upon movers only, i.e., those individuals who actually switch groups as part of the reallocation.[7] This is in contrast to the external effect which operates on stayers.

First consider the targeting effect, if X and S co-vary under the status quo, then implementing (7) will alter the conditional distribution of X for high and low type individuals. For example, if under the status quo $Cov(X,S) > 0$, then a segregation-increasing reallocation will require high type agents to

[7] Since these effects only operate upon movers, they would be, in principle, priced appropriately by a competitive market for group membership.

move into groups with, not only above average initial levels of S, but also with larger levels of X.[8] To the extent that high and low types are differentially affected by changes in X, this will alter expected outcomes even in the absence of social spillovers.

Consider once again the classroom setting. Krueger and Whitmore (2002) present evidence that black elementary school students benefit from class size reductions more than white students (in terms of increased test scores). If black students tend to be located in larger classrooms than whites then, even in the absence of peer group effects, integrating classrooms will raise average achievement by improving input targeting. Such reallocations reduce average class size for blacks while raising it for whites, but the outcome benefits of the former effect outweigh the outcome costs of the latter. If, on the other hand, class size tends to be smaller for blacks, then integrating classrooms will, in the absence of some compensating peer group effect, lower average achievement. In general, moves toward greater segregation or integration may improve or worsen the targeting of other inputs (such as teacher quality or class size).

The sign of $Cov\left(p\left(S,X\right),S\right)$ also reflects the direction of the *private peer effect.* If the benefits of improved peer quality for high type movers entering groups with an initially above average fraction of high types exceed the costs for low type movers leaving such groups, then the reallocation will tend to raise average achievement of movers. Observe that the private peer effect will be zero when outcomes are separable in own and peer types (as is often assumed in empirical work), positive when they are complementary and negative when they are substitutable.

The better understand the distinction between the two effects it is helpful to consider the additively separable case where $m_H\left(s,x\right)=m_H\left(s\right)+g_H\left(x\right)$ and analogously for $m_L\left(s,x\right)$. Letting $p\left(s\right)=m_H\left(s\right)-m_L\left(s\right)$ and $t\left(x\right)=g_H\left(x\right)-g_L\left(x\right)$, in this case we have

$$Cov\left(p\left(S,X\right),S\right)=Cov\left(p\left(S\right),S\right)+Cov\left(t\left(X\right),S\right),$$

which correspond directly to the private peer and targeting effects.

In order to decompose $Cov\left(p\left(S,X\right),X\right)$ into its private peer and targeting effect components in the general case it is helpful to begin with the observation that under random assignment to groups we would have independence of S and X (i.e., $f_{S,X}\left(s,x\right)=f_S\left(s\right)f_X\left(s\right)$). In such settings implementing (7) would not alter the conditional distribution of X given S and $Cov\left(p\left(S,X\right),S\right)$ would reflect solely the private peer effect, since independence of S and X under the status quo ensures that the targeting effect is zero.

Redefine $p\left(s\right)$ to equal $\mathbb{E}_X\left[p\left(s,X\right)\right]$. The sign of the estimand

$$\mathbb{E}_S\left[p\left(S\right)\left(S-\mu_T\right)\right]=Cov\left(p\left(S\right),S\right)$$

reflects only the direction of private peer effects on average outcomes. By taking expectations over the product of marginals we have artificially imposed independence of S and X. This gives a clean interpretation to $Cov\left(p\left(S\right),S\right)$ as the (non-spillover) effect associated with implementing (7) that would have occurred if S and X were independent under the status quo.

[8] Let $\mathbb{E}^*[S|X=x]=\alpha+\beta x$ denote the (mean squared error minimizing) linear predictor (LP) of an individual's peer group composition given other inputs X. After the reallocation we have $\mathbb{E}^*[S^*|X=x]=\left(\alpha-\lambda\beta\mu_X\right)+\beta\left(1+\lambda\right)x$.

We define the targeting effect of implementing (7) as the difference

$$\mathbb{E}\left[p\left(S,X\right)\left(S-\mu_T\right)\right]=\mathbb{E}\left[p\left(S\right)\left(S-\mu_T\right)\right]+\mathbb{E}\left[t\left(S,X\right)\left(S-\mu_T\right)\right],$$

where $t\left(s,x\right)=p\left(s,x\right)-p\left(s\right)$.

To summarize we can decompose the average outcome effects of implementing (7) into three parts:

$$Cov\left(m'\left(S,X\right),S\right)=\underbrace{Cov\left(p\left(S\right),S\right)}_{\text{Private Peer Effect}}+\underbrace{Cov\left(t\left(S,X\right),S\right)}_{\text{Targeting Effect}}+\underbrace{Cov\left(e\left(S,X\right),S\right)}_{\text{Spillover Peer Effect}}.$$

It is also interesting to measure the effects of implementing (7) on inter-type inequality. The change in the expected outcomes for high and low type individuals associated with implementing reallocation (7), in a group with $S=s$ and $X=x$, are given, respectively, by

$$m_H\left(s+\lambda\left(s-\mu_T\right),x\right)-m_H\left(s,x\right),\qquad m_L\left(s+\lambda\left(s-\mu_T\right),x\right)-m_L\left(s,x\right).$$

The difference in average outcomes between high and low types therefore grows (shrinks) under the local reallocation defined by (7) if

$$\mathbb{E}\left[m'_H\left(S,X\right)\left(S-\mu_T\right)|T=H\right]-\mathbb{E}\left[m'_L\left(S,X\right)\left(S-\mu_T\right)|T=L\right] \tag{9}$$

is positive (negative).

Locally increasing segregation requires reassigning high type individuals from groups with below to above average fractions of high types. Each reassigned high type is switched with a low type. If (9) is positive then gains to high types from this pattern of reallocation will exceed those of low types. This is because high types in groups *receiving* an additional high type tend to gain more from the change than low types in such groups *and* high types in groups *sending* a high type tend to lose less from the change then low types in such groups. For small λ, the product of λ and (9) gives the approximate change in the outcome gap.

2.3 Connections to the theory of sorting in the presence of social spillovers

It is illuminating to relate (8) and (9) to the theoretical literature on sorting in the presence of social spillovers. In order to make these connections we assume that an individual's willingness-to-pay (to be in a specific group) is linear function of own outcomes (that does not vary by type). We also assume that both types of individuals receive equal weight in the social planner's objective function. The social planner's objective is to maximize average outcomes.[9]

We can now interpret a test of $Cov\left(m'\left(S,X\right),S\right)=0$ as a test of whether the status quo allocation is locally efficient. Consider a social planner with full knowledge of $m\left(s,t,x\right)$ and its derivatives with respect to s, but without knowledge of individual potential outcomes. For simplicity assume that the social planner's optimal allocation of individuals to groups does not involve corner solutions (i.e.,

[9] It would be straightforward, as well as fruitful, to extend the discussion that follows to allow for alternative objective functions. For example a Rawlsian social planner would seek to maximize expected outcomes for low types (c.f., Piketty 2004).

$S_c \in (0,1)$ for $c = 1, \dots, C$). Under these conditions the social planner chooses $S_1, \dots, S_C$ such that the marginal effect of changes in S_c on expected average outcomes is equated across all groups:

$$m'(S_1, X_1) = \ldots = m'(S_C, X_C).$$

Deviations of $Cov(m'(S,X), S)$ from zero thus provide evidence against the null of an efficient status quo allocation.[10] The sign of this covariance indicates which direction – more or less segregation – leads to increases in average outcomes.

In the presence of social spillovers the efficient allocation of individuals to groups may deviate from the pattern of segregation observed in a competitive equilibrium. Perfect competition equalizes high and low types' willingness-to-pay for a move across groups. If individual willingness-to-pay is an increasing linear function of own outcomes then competition will also equalize the expected outcome gap between high and low types across groups. If, for example, the expected outcome gain associated with a move from group A to group B for high types' exceeds that of low types, then high types in A will 'buy out' low types in B; such behavior is inconsistent with a competitive equilibrium. Under *laissez faire* we therefore have

$$p(S_1, X_1) = \ldots = p(S_N, X_N),$$

and hence $Cov(p(S,X), S) = 0$ when the status quo allocation is also a competitive equilibrium.

When $Cov(p(S,X), S) \neq 0$ there are private incentives for continued sorting, i.e., for individuals to trade places across groups (although individuals may be unable to act upon such incentives in the presence of credit constraints or other market inefficiencies). If $Cov(p(S,X), S) > 0$, then high types in groups with few high type peers ($s < \mu_T$) will buy out low types in groups with many such peers ($s > \mu_T$), increasing segregation. This pattern of switching is feasible since the private outcome gains for high types involved in such switches will exceed the outcome losses for their low type trading partners. If $Cov(p(S,X), S) < 0$ the opposite occurs, generating more integrated groups.

If the status quo allocation is administratively determined, or individuals are otherwise constrained in their choice of social groups (e.g., by imperfect credit markets), then it may not correspond to a *laissez-faire* equilibrium. In such cases it is interesting to predict whether a policy which implies a greater role for market forces in the allocation of individuals to groups would increase or decrease segregation. As mentioned in the introduction, implementing certain policies, such as 'school choice', would presumably increase the role of market forces in assigning individuals to groups. The sign of $Cov(p(S,X), S)$ provides information on the likely effect of such policies on segregation. If $Cov(p(S,X), S) < 0$, then such policies might enable ghettoized low types to move into groups with more high type peers. If $Cov(p(S,X), S) > 0$, such policies might simply enable high types to more effectively segregate themselves from low types.

The sign of $Cov(p(S,X), S)$ indicates whether there are incentives for continued sorting and hence whether the private marginal benefit of moving into a group with many high types is greater for high versus low type individuals. The net social impact of such switches on average outcomes, however, depends upon the sign of the *sum* of $Cov(p(S,X), S)$, the private costs and benefits reaped by movers,

[10] For a more complete discussion of this efficiency condition see de Bartolome (1990), Benabou (1993) and Becker and Murphy (2000), all of whom analyze the social planner's problem in settings very similar to our own.

and $Cov(e(S,X),S)$, the social spillover imposed by movers upon stayers. When contemplating a move across groups individuals do not internalize the positive and negative effects of their actions on their peers. Consequently there may be too much or too little segregation in equilibrium.

To better understand the potential divergence between the competitive equilibrium and the efficient allocation, as well as to highlight connections to theory, it is useful to consider once again the additively separable case where $m_H(s,x) = m_H(s) + g_H(x)$ and analogously for $m_L(s,x)$. Recalling that $p(s) = m_H(s) - m_L(s)$, $t(x) = g_H(x) - g_L(x)$ and $e(s) = sm'_H(s) + (1-s)m'_L(s)$ in this case, we have

$$\frac{Cov(m'(S,X),S)}{Var(S)} = \frac{Cov(p(S),S)}{Var(S)} + \frac{Cov(t(S),S)}{Var(S)} + \frac{Cov(e(S),S)}{Var(S)}.$$

Where, using Lemma 2 in Appendix A ,

$$\begin{aligned} \frac{Cov(p(S),S)}{Var(S)} &= \mathbb{E}[\omega(S)\left(m'_H(S) - m'_L(S)\right)] \\ \frac{Cov(t(S),S)}{Var(S)} &= \mathbb{E}\left[\omega(S)\left(\frac{\partial E[g_H(X)|S]}{\partial S} - \frac{\partial E[g_L(X)|S]}{\partial S}\right)\right] \\ \frac{Cov(e(S),S)}{Var(S)} &= \mathbb{E}\left[\omega(S)\left(m'_H(S) - m'_L(S) + Sm''_H(S) + (1-S)m''_L(S)\right)\right]. \end{aligned} \tag{10}$$

where $\mathbb{E}_S[\omega(S)] = 1$. The precise form of $\omega(s)$ is given in the Appendix; it emphasizes values S close to μ_T.

Equation (10) gives a weighted average derivative representation to the slope coefficient of the (mean squared error minimizing) linear predictor (LP) of $m'(S,X)$ given S. A test of whether the first component of this expression – $Var(S)^{-1}Cov(p(S),S)$ – equals zero is a test of the null of separability between own and peer type (i.e., of whether the marginal benefit of a change in peer composition does not depend on own type). Under the alternative $Var(S)^{-1}Cov(p(S),S)$ equals a weighted average of 'net complementarity' or 'net substitutability' between own and peer type. It is the presence of net complementarity, of course, that drives stratification in theoretical models of sorting (e.g., Benabou 1993, p. 626).

The sign of $Var(S)^{-1}Cov(t(S),S)$ captures the influence of inequities in levels of X across groups on private incentives to sort. de Bartolome (1990) has an extensive discussion of this channel for sorting in his model. The channel is also formally similar to the impact of differences in amenities across communities on sorting in Becker and Murphy (2000). Finally, as discussed previously, the sign of the final term determines whether a local increase in segregation will lower or raise average spillover strength.

Note also that $Cov(m'(S,X),S)/Var(S)$ equals a weighted average of the second derivatives of expected average outcomes with respect to group composition. It therefore provides information on the degree to which $m(s,x)$ is convex in s. The effect of implementing (7) on average outcomes depends on the sign of a weighted average of $m''(s,x)$, i.e., upon convexity of $m(s,x)$ in s local to its expectation. The appropriate weighted average for identifying the direction of the reallocation's effect on average outcomes is given by (10), c.f., the Corollary to Lemma 2 in Appendix A.

If $Cov(p(S),S) + Cov(t(S),S) + Cov(e(S),S)$ is negative there are decreasing social returns to stratification by type. However if $Cov(p(S),S) + Cov(t(S),S)$ is positive there may be private incen-

Table 1: Effects of moving toward market allocation of individuals to groups

Panel A: Signs of covariances	(1)	(2)	(3)	(4)
$Cov\left(p\left(S,X\right),S\right)$	$+$	$+$	$-$	$-$
$Cov\left(m'\left(S,X\right),S\right)$	$+$	$-$	$-$	$+$
Panel B: Direction of shift toward laissez-faire on				
Segregation	$+$	$+$	$-$	$-$
Average outcomes	$+$	$-$	$+$	$-$

NOTES: For policy shifts that represent a shift toward *laissez-faire,* this table reports the sign of the effect on segregation and average outcomes.

tives for continued, but socially inefficient, sorting. To the extent that we can view a policy as a shift toward a *laissez-faire* allocation, the signs of $Cov\left(p\left(S,X\right),S\right)$ and $Cov\left(m'\left(S,X\right),S\right)$ provide valuable information on effects on such a policy on segregation and average outcomes respectively. Table 1 summarizes the main relationships. Panel A of the table gives the four possible sign combinations for $Cov\left(p\left(S,X\right),S\right)$ and $Cov\left(m'\left(S,X\right),S\right)$. Panel B gives the direction of the impact of a move toward *laissez-faire* on segregation and average outcomes.

Consider column 1, if under the status quo $Cov\left(p\left(S,X\right),S\right)$ is positive, then a move toward *laissez-faire* will induce more sorting. Such sorting is efficient in this case – in the sense that it raises average outcomes – since $Cov\left(m'\left(S,X\right),S\right)$ is also positive. In column 2, private incentives are not aligned with what is socially desirable and a shift toward *laissez-faire* results in an inefficient increase in segregation. This is a case emphasized by much of the theoretical work on sorting in the presence of social externalities (e.g., Benabou 1993, Piketty 2000). In column 3 private incentives are again aligned with what is socially desirable: segregation declines while average outcomes rise. Finally, in column 4, a shift toward *laissez-faire* reduces segregation but also lowers average outcomes.

3 Estimands in the context of standard parametric models

This section relates the estimands outlined in the previous section to common parametric models of social interactions. The majority of empirical work on social externalities can be interpreted as assuming that $m\left(s,t,x\right)$ equals

$$m\left(s,t,x\right)=\alpha_{t}+\beta s+x'\gamma. \tag{11}$$

Equation (11) can be interpreted as the reduced form of the 'linear-in-means' model of social interactions (c.f., Equation (5) in Manski (1993, p. 534)). This model is the workhorse of empirical work on social interactions.[11] From (11) we have

$$p\left(s,x\right)=\left(\alpha_{H}-\alpha_{L}\right),\quad e\left(s,x\right)=\beta,$$

and hence $Cov\left(m'\left(S,X\right),S\right)=Cov\left(p\left(S,X\right),S\right)=Cov\left(e\left(S,X\right),S\right)=0$. The linear-in-means model rules out an equity versus efficiency trade-offs by construction. The status quo allocation generates the

[11] Mayer and Jencks (1989), Solon (1999), Ginther, Haveman and Wolfe (2000) provide alternative surveys of this literature.

same average outcomes as all other feasible allocations. If willingness-to-pay is linear in outcomes, then the model also fails to provide a rational for sorting by types. Allocations do affect inequality by type, with (9) equaling

$$\beta\left(\mathbb{E}\left[S|T=H\right]-\mathbb{E}\left[S|T=L\right]\right).$$

As noted by Graham (2005) and others (11) is a very restrictive model of social interactions. A simple generalization, with substantially richer implications, but less often used in applied work[12] is

$$m\left(s,t,x\right)=\alpha_t+\beta_t s+x'\gamma_t. \tag{12}$$

Equation (12) allows the effect of peer composition and group characteristics to vary by type. This feature ensures that (12) is sufficiently general to illustrate the full range of issues discussed in Section 2. We have

$$\begin{aligned} p\left(s,x\right) &= \left(\alpha_H-\alpha_L\right)+\left(\beta_H-\beta_L\right)s+x'\left(\gamma_H-\gamma_L\right), \\ e\left(s,x\right) &= s\cdot\beta_H+\left(1-s\right)\cdot\beta_L. \end{aligned}$$

The effect on average outcomes from implementing (8) in this model is therefore given by

$$Cov\left(m'\left(S,X\right),S\right)= \underbrace{\left(\beta_H-\beta_L\right)Var\left(S\right)}_{\text{Private Peer Effect}}+\underbrace{\left(\gamma_H-\gamma_L\right)'Cov\left(X,S\right)}_{\text{Targeting Effect}}+\underbrace{\left(\beta_H-\beta_L\right)Var\left(S\right)}_{\text{Spillover Peer Effect}},$$

while the effect on inter-type inequality is given by

$$\beta_H\mathbb{E}\left[S|T=H\right]-\beta_H\mathbb{E}\left[S|T=L\right].$$

4 Identification

This section provides a very brief discussion of identification. We work with a standard 'selection on observables' type assumption.

Condition 1 (Selection on Observables) $Y\left(s\right)\perp S|X$ *for* $s\in[0,1]$

Condition 1 implies that within subpopulations defined by X, assignment to groups, and hence group composition, is 'as if' randomly assigned. Under Condition 1 we have

$$\mathbb{E}\left[Y|S=s,X=x\right]=\mathbb{E}\left[Y\left(s\right)|X=x\right]=m\left(s,x\right),$$

and

$$\mathbb{E}\left[Y|S=s,X=x,T=t\right]=\mathbb{E}\left[Y\left(s\right)|X=x,T=t\right]=m\left(s,t,x\right).$$

[12] Equation (12) is most often used in research which studies the effects of school desegregation, where differential effects on black and white achievement are of considerable interest. Schofield (1995) provides a comprehensive review of research in education and sociology on this topic. Guryan (2004) and Card and Rothstein (2005) provide recent examples of economic research in this area. Both of these papers work with a model similar to (12).

Thus $m_H(s,x)$, $m_L(s,x)$ and $m(s,x)$ are identified by the appropriate sample means. Identification of the derivatives of these functions with respect to s also follows under standard conditions.

Condition 1 is often a strong, and in some settings, an incredible assumption. Its appropriateness for identifying models of social spillovers in settings where individuals exercise substantial discretion over group membership is debatable. However, we believe that in some settings it can be very compelling, and in others, a reasonable point of departure.

An obvious setting where Condition 1 is likely to be satisfied is when the data are generated by explicit experimentation. For example, the Moving to Opportunity (MTO) demonstration experiment, analyzed by Kling, Liebman and Katz (2005) among others, can be viewed as generating exogenous – in the sense of Condition 1 – variation in the fraction of one's neighbors below the poverty line.[13] The assumption is also plausible in other settings. For example, variation in the gender composition across adjacent cohorts within the same elementary school is arguably idiosyncratic (c.f., Hoxby 2002). More generally the 'exchangeability' arguments used by Altonji and Matzkin (2005) will be useful in social spillover applications.

With $m_H(s,x)$, $m_L(s,x)$ and their derivatives in hand, the estimands defined in Section 2 are identified by standard method-of-moments arguments. Let $W = (S, X')'$ and $\gamma(w) = (m_H(s,x), m_L(s,x), \mu_T)'$, then $\beta_0 = Cov(p(S,X), S)$, for example, is estimable using the moment function

$$\psi(Z, \beta, \gamma(W)) = (\gamma_1(W) - \gamma_2(W))(S - \gamma_3(W)) - \beta. \tag{13}$$

In practice, $m_H(s,x)$, $m_L(s,x)$ and μ_T in (13) would be replaced by nonparametric estimates leading to a semiparametric two-step estimator for β_0. Characterizing the large-sample properties of such an estimator is non-trivial but can be done using the results of Newey (1994) and Newey and McFadden (1994). This is a topic of ongoing research.

5 Conclusion

This paper has outlined a new framework for measuring the strength, and characterizing the nature, of social spillovers or peer group effects. The majority of previous empirical work on social spillovers, at best, provides a consistent test of the null of no social spillovers (e.g., Graham 2005). In defining our proposed estimands we have been guided by the applied public finance theory on sorting across neighborhoods. As a result our framework provides a basis for both testing the null of no social spillovers as well as characterizing the nature of any such spillovers in a policy-relevant way. In particular, we provide approaches for characterizing the equity and efficiency consequences of local departures from the status quo allocation of individuals to groups.

In ongoing work we are developing feasible estimators for the objects defined in Section 2 as well as deriving their large sample properties. We are also working on an empirical application relating student achievement in elementary school to the gender composition of one's classmates. In the longer-run, it would be useful to provide identification results under weaker conditions.

[13] All MTO participants were eligible for public housing under standard conditions. Some vouchers required families to move into low poverty neighborhoods, while others did not. Since all voucher recipients were themselves poor the experiment only identifies – in the notation of this paper – $m_L(s,x)$. A parallel experiment, designed to generate variation in neighborhood composition among the non-poor, could identify $m_H(s,x)$.

Appendices

A Derivations and proofs

Lemma 2 *For S, a continuous random variable, with (i) finite support $S \in [a,b]$, (ii) cumulative distribution function $F(s)$, and (iii) $g(\cdot)$ a continuously differentiable function on the support of S:*

1. *The slope coefficient of the (mean squared error (MSE) minimizing) linear predictor (LP) of $g(S)$ given S has a weighted average derivative representation of*

$$\beta = \frac{Cov(g(S), S)}{Var(S)} = \mathbb{E}\left[\omega(S)\frac{\partial g(S)}{\partial s}\right],$$

where

$$\omega(t) = \frac{1}{dF(t)} \frac{\mathbb{E}[S-\mu_S|S \geq t](1-F(t))}{\int_{v=a}^{v=b}\mathbb{E}[S-\mu_S|S \geq v](1-F(v))\,dv}, \quad \mathbb{E}[\omega(S)] = 1,$$

and

2. *β gives maximum weight to values of $\frac{\partial g(S)}{\partial s}$ for S close to its mean, μ_S, and minimum weight when S is near the boundaries of its support.*

Proof. The proof for the first result is similar to that of Lemma 5 of Angrist, Imbens and Graddy (2000).[14] The second result of the Lemma, i.e., the precise characterization of the weighting process follows from a simple integration by parts argument. Observe that $g(S) - g(a) = \int_{t=a}^{t=S}\frac{\partial g(t)}{\partial t}dt$ and that $E[g(a)(S-\mu_S)] = 0$ for $\mu_S = E[S]$. Under weak conditions we therefore have

$$\begin{aligned}
Cov(g(S), S) &= \mathbb{E}[g(S)(S-\mu_S)] \\
&= \mathbb{E}\left[\int_{t=a}^{t=S}\frac{\partial g(t)}{\partial t}(S-\mu_S)\,dt\right] \\
&= \mathbb{E}\left[\int_{t=a}^{t=b}\frac{\partial g(t)}{\partial t}(S \geq t)(S-\mu_S)\,dt\right] \\
&= \int_{t=a}^{t=b}\frac{\partial g(t)}{\partial t}\mathbb{E}[(S \geq t)(S-\mu_S)]\,dt \\
&= \int_{t=a}^{t=b}\frac{\partial g(t)}{\partial t}\mathbb{E}[S-\mu_S|S \geq t](1-F(t))\,dt.
\end{aligned}$$

The variance of S can be written as

$$\begin{aligned}
Var(S) &= \mathbb{E}\left[S(S-\mu_S)'\right] \\
&= \mathbb{E}\left[\int_{v=a}^{v=S} 1(S-\mu_S)\,dv\right] \\
&= \int_{v=a}^{v=b}\mathbb{E}[S-\mu_S|S \geq v](1-F(v))\,dv.
\end{aligned}$$

[14] Sarychev (2001) and Inoue (2002) also provide weighted average derivative represention of β.

The first result follows for $\omega(t)$ as given in the Lemma. To show the second result, that the weighted average derivative representation of β gives the most emphasis to values of $\frac{\partial g(S)}{\partial s}$ for S close to its mean, begin by noting that

$$\mathbb{E}\left[\frac{\partial g(S)}{\partial s}\omega(S)\right] = \frac{\int_{t=a}^{t=b}\frac{\partial g(s)}{\partial s}\mathbb{E}\left[S-\mu_S|S\geq t\right](1-F(t))\,dt}{\int_{v=a}^{v=b}\mathbb{E}\left[S-\mu_S|S\geq v\right](1-F(v))\,dv}.$$

Therefore the size of the weight on $\frac{\partial g(t)}{\partial t}$ is proportional to

$$\mathbb{E}\left[S-\mu_S|S\geq t\right](1-F(t)).$$

Integration by parts (with $u = 1 - F(s)$ and $v = s$) gives

$$\begin{aligned}\int_t^b \left[1-F(s)\right]\, ds &= \left\{1-F(s)s\right\}|_t^b - \int_t^b s\ dF(s) \qquad (14)\\ &= \left[1-F(t)\right]t - \int_t^b s\ dF(s).\end{aligned}$$

We then write

$$\begin{aligned}\frac{\partial}{\partial t}\left\{E\left[S-\mu_S|S\geq t\right]\left[1-F(t)\right]\right\} &= \frac{\partial}{\partial t}\int_t^b s\ dF(s) - \frac{\partial}{\partial t}\left[1-F(t)\right]\mu_S\\ &= \frac{\partial}{\partial t}\int_t^b s\ dF(s) + \mu_S\cdot dF(t)\end{aligned}$$

Using (14) to substitute for $\int_t^b s\ dF(s)$ gives

$$\begin{aligned}\frac{\partial}{\partial t}\left\{E\left[S-\mu_S|S\geq t\right]\left[1-F(t)\right]\right\} &= \frac{\partial}{\partial t}\left\{\left[1-F(t)\right]t - \int_t^b \left[1-F(s)\right]ds\right\}\\ &\quad +\mu_S\cdot dF(t)\\ &= \left[1-F(t)\right] - \frac{\partial}{\partial t}\int_t^b \left[1-F(s)\right]ds\\ &\quad -(t-\mu_S)\cdot dF(t)\\ &= \left[1-F(t)\right]-\left[1-F(t)\right]-(t-\mu_S)\cdot dF(t)\\ &= -(t-\mu_S)\cdot dF(t).\end{aligned}$$

This gives $\frac{\partial}{\partial t}\left\{E\left[S-\mu_S|S\geq t\right]\left[1-F(t)\right]\right\} = 0$ at $t = \mu_S$. This derivative is negative for $t > \mu_S$ and positive for $t < \mu_S$, hence it attains a maximum at $t = \mu_S$ and its minimum at the boundaries of the support of S. *Q.E.D.* ■

Corollary 3 *Let $m(t) = \mathbb{E}[Y(t)]$ equal the average potential outcome associated with the continuously valued treatment $t \in [a, b]$. Treatment reallocations of the form $T^* = T + \lambda(T - \mu_T)$ with $\lambda \to 0$ from above and $\mu_T = \mathbb{E}[T]$ will raise average outcomes if and only if $m(t)$ is 'locally convex' in the sense that*

$$\mathbb{E}\left[\omega(T)\frac{\partial^2 m(T)}{\partial t\partial t}\right]$$

is positive.

Proof. The change in average outcomes is given by

$$\mathbb{E}\left[m\left(T^{*}\right)-m\left(T\right)\right]=\mathbb{E}\left[m\left(T+\lambda\left(T-\mu_{T}\right)\right)-m\left(T\right)\right],$$

differentiating through the expectation operator with respect to λ and evaluating at $\lambda = 0$ gives the direction of the reallocation's effect on average outcomes:

$$\mathbb{E}\left[m'\left(T\right)\left(T-\mu_{T}\right)\right]=Cov\left(m'\left(T\right),T\right).$$

The claim then follows from Lemma 2. *Q.E.D.* ■

Remark 4 *A least squares regression of $m'(T)$ on T identifies the direction of the effect of the reallocation on average outcomes.*

References

Altonji, Joseph G. and Rosa L. Matzkin (2005). "Cross section and panel data estimators for nonseparable models with endogenous regressors," *Econometrica* 73 (4): 1053 - 1102.

Angrist, Joshua D., Kathryn Graddy and Guido W. Imbens. (2000). "The interpretation of instrumental variables estimators in simultaneous equations models with an application to the demand for fish," *Review of Economics Studies* 67 (3): 499 - 527.

Angrist, Joshua D. and Kevin Lang. (2004). "Does school integration generate peer effects? Evidence from Boston's Metco Program," *American Economic Review* 94 (5): 1613 - 1634.

Bayer, Patrick, Fernando Ferreira and Robert McMillan. (2003). "A unified framework for measuring preferences for schools and neighborhoods," *Mimeo*, Yale University.

Bayer, Patrick, Randi Pintoff and David E. Pozen. (2004). "Building criminal capital behind bars: peer effects in juvenile corrections," *Mimeo*, Yale University.

Becker, Gary S. and Kevin M. Murphy. (2000). *Social Economics: Market Behavior in a Social Environment.* Cambridge, MA: Harvard University Press.

Benabou, Roland. (1993). "Workings of a city: location, education, and production," *Quarterly Journal of Economics* 108 (3): 619 - 652.

Brock, William A. and Steven N. Durlauf. (2001). "Interactions-based Models," *Handbook of Econometrics* 5: 3297 - 3380 (J. Heckman & E. Leamer, Eds.). Amsterdam: North-Holland.

Card, David and Jesse Rothstein. (2005). "Racial segregation and the black-white test score gap," *Mimeo.*

Coleman, James S. et al. (1966). *Equality of Educational Opportunity.* Washington: U.S. Department of Health, Education and Welfare.

de Bartolome, Charles A. M. (1990). “Equilibrium and inefficiency in a community model with peer group effects,” *Journal of Political Economy* 98 (1): 110 - 133.

Ekeland, Ivar, James J. Heckman and Lars Nesheim. (2004). “Identification and estimation of hedonic models,” *Journal of Political Economy* 112 (1): S60 - S108.

Epple, Dennis. (1987). “Hedonic prices and implicit markets: estimating demand and supply functions for differentiated products,” *Journal of Political Economy* 95 (1): 59 - 80.

Epple, Dennis and Richard Romano. (1998). “Competition between private and public schools, vouchers, and peer-group effects,” *American Economic Review* 88 (1): 33 - 62.

Epple, Dennis, Elizabeth Newlon and Richard Romano. (2002). “Ability tracking, school competition, and the distribution of educational benefits,” *Journal of Public Economics* 83 (1): 1 - 48.

Epple, Dennis and Holger Sieg. (1999). “Estimating equilibrium models of local jurisdictions,” *Journal of Political Economy* 107 (4): 645 - 681.

Fernandez, Raquel. (2003). “Sorting, education and inequality,” in *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress* 2: 1 - 40 (M. Dewatripont, L. P. Hansen, S.J. Turnovsky, Eds.). Cambridge: Cambridge University Press.

Figlio, David N. and Marianne E. Page. (2002). “School choice and the distributional effects of ability tracking: does separation increase inequality?” *Journal of Urban Economics* 51 (3): 497 - 514.

Ginther, Donna, Robert Haveman and Barbara Wolfe. (2001). “Neighborhood attributes as determinants of children’s outcomes,” *Journal of Human Resources* 35 (4): 603 - 642.

Glaeser, Edward and José A. Scheinkman. (2003). “Nonmarket interactions,” *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress* 1: 339 - 369 (M. Dewatripont, L. P. Hansen, S.J. Turnovsky, Eds.). Cambridge: Cambridge University Press.

Graham, Bryan S. (2005). “Identifying social interactions through excess variance contrasts,” *Mimeo.*

Guryan, Jonathan. (2004). “Desegregation and black dropout rates,” *American Economic Review* 94 (4): 919 - 943.

Hoxby, Caroline M. (2002). “The power of peers: how does the makeup of a classroom influence achievement?” *Education Next* 2 (2): 57 - 63.

Inoue, Atsushi. (2002). “Identifying the sign of the slope of a monotonic function via OLS,” *Economic Letters* 73 (3): 419 - 424.

Kling, Jeffrey R, Jeffrey B. Liebman and Lawrence F. Katz. (2005). “Experimental analysis of neighborhood effects,” *NBER Working Paper No. 11577.*

Krueger, Alan and Diane Whitmore. (2002). "Would smaller classes help close the black-white achievement gap?" *Bridging the Achievement Gap*: 11 - 46 (J.E. Chubb & T. Loveless, Eds.). Washington D.C.: Brookings Institution Press.

Lazear, Edward P. (2001). "Educational production," *Quarterly Journal of Economics* 116 (3): 777 - 803.

Manski, Charles F. (1993). "Identification of endogenous social effects: the reflection problem," *Review of Economic Studies* 60 (3): 531 - 542.

Mayer, Susan E. and Christopher Jencks. (1989). "Growing up in poor neighborhoods: how much does it matter," *Science* 243 (4897): 1441 - 1445.

Newey, Whitney K. (1994). "The asymptotic variance of semiparametric estimators," *Econometrica* 62 (6): 1349 - 1382.

Newey, Whitney K. and Daniel McFadden. (1994). "Large sample estimation and hypothesis testing," *Handbook of Econometrics 4*: 2111 - 2245 (R.F. Engle & D.L. McFadden). Amsterdam: North Holland.

Piketty, Thomas. (2000). "Theories of persistent inequality and intergenerational mobility," *Handbook of Income Distribution* 1: 430 - 476 (A. Atkinson & F. Bourguignon). Amsterdam: North Holland.

Piketty, Thomas. (2004). "L'Impact de la taille des classes et de la ségrégation sociale sur la réussite scolaire dans les écoles françaises : une estimation à partir du panel primaire 1997," *Mimeo.*

Sarychev, Andre (2001). "Neighborhood sorting, linear regressions and human capital inequality," *Mimeo.*

Schofield, Janet Ward. (1995). "Review of research on school desegregation's impact on elementary and secondary school students," *Handbook of Research on Multicultural Education*: 597 - 616 (J.A. Banks & C.A.M. Banks). New York: Macmillan Publishing USA.

Solon, Gary. (1999). "Intergenerational mobility in the labor market," *Handbook of Labor Economics* 3: 1761 - 1800 (O. Ashenfelter & D. Card, Eds.). Amsterdam: North-Holland.